

# **The Coarticulation of Bi-Syllabic Words in Chinese**

**Maolin Wang, Wei Xiong**

College of Chinese Language and Culture, Jinan University, Guangzhou, China

---

**Abstract:-** In this study, trans-consonantal vowel-to-vowel coarticulation in Chinese is investigated. The target words are in the form of 'bVl.ba', and the subjects are eight native speakers of standard Chinese. Vowel formants are examined at the onset, middle and offset points of the target vowel. Results show that trans-segmental coarticulation exists in Chinese, especially for the second formant value of the vowel, and in Chinese, coarticulatory effect does not extend to the offset point of the vowel.

**Keywords:-** Speech, coarticulation, vowel, formant, syllable

---

## **I. INTRODUCTION**

This study deals with the extent of trans-consonantal vowel-to-vowel coarticulation in bi-syllabic words in Chinese. Coarticulation refers to the fact that speech sounds are not produced as isolated gestures, but are superimposed in a complex, context-dependent fashion. While it is clear that all natural speech is coarticulated, the magnitude and temporal extent of coarticulation in different contexts has been difficult to explain. The need to adequately describe the varieties of coarticulation has been a key factor driving the development of competing models of speech production and also perception [1].

It is shown from researches on speech production that there are systematic differences between languages in the spatiotemporal characteristics of coarticulation [2]. Öhman [3] compared the coarticulatory effects in three languages, and found that the  $F_2$  values of target vowels varied more in English and Swedish than in Russian, due to vowel context. He attributed the coarticulatory differences to the languages' consonant systems, arguing that the requirements on the tongue body imposed by contrastive palatalization in Russian restricted trans-consonantal coarticulation.

Now that the effect of transconsonantal vowel-to-vowel coarticulation is diversified among languages, researchers have sought to understand how different factors influence this effect, and it is shown that a language's system of vowel contrasts may influence V-to-V effect. Beddor et al. [4] conducted three experiments to test the hypothesis that V-to-V coarticulatory organization differs in Shona and English. An acoustic study of Shona and English trisyllables shows that the two languages differ in the coarticulatory effect of stressed and unstressed vowels on each other, and the relation between the production and perception data suggests that listeners are attuned to native-language coarticulatory patterns.

Research results show that languages with larger vowel systems tend to exhibit weaker V-to-V coarticulatory effects than those with smaller systems. Weaker effects have been shown for English compared to the much smaller five-vowel systems of Shona, Swahili [5]. Cho [6] examined how the degree of vowel-to-vowel coarticulation varies as a function of prosodic factors, and results show that vowels in prosodically stronger locations are coarticulated less with neighboring vowels, but do not exert a stronger influence on the articulation of neighboring vowels. An examination of the relationship between coarticulation and duration reveals that accent-induced coarticulatory variation cannot be attributed to a duration factor, and some of the data with respect to boundary effects may be accounted for by the duration factor. He proposed that prosodically conditioned V-to-V coarticulatory reduction is another type of strengthening that occurs in prosodically strong locations. The prosodically driven coarticulatory pattern can be taken as part of the phonetic signaling of the hierarchically nested structure of prosody.

In regard to the extent of coarticulation, studies on various languages have shown vowel-to-vowel coarticulatory effects not only in transitions, but extending into the steady-state period of the transconsonantal vowel both in palatographic data [7-8] and in acoustic data [9-10]. While there is ample evidence of the existence of vowel-to-vowel coarticulatory effects, factors have been cited which affect the extent of these effects. For instance, these effects may be constrained by intervocalic palatals and velars, whose production requires use of the tongue body in conflict with the production of vowels, thereby restricting vowel-to-vowel coarticulation [11].

It is indicated in early studies that V-to-V coarticulation might be a relatively local phenomenon [12], subsequent work has shown that this is not always the case. Instances of long-distance coarticulation, involving effects crossing two or more intervening segments have been found [13-14]. Magen [15] analyzed [bVbābVb] sequences produced by four English speakers and found evidence of coarticulatory effects between the first and final vowel, meaning that such effects can cross foot boundaries and multiple syllable boundaries. More

recently, Grosvald [16] studied long-distance vowel-to-vowel coarticulation in English, and found that anticipatory V-to-V coarticulation can occur over at least three vowels' distance in natural discourse, and that even such long-distance effects can be perceived by some listeners.

There has been some research work on the coarticulation of segments in Chinese. Wu and Sun [17] analyzed the acoustic coarticulatory patterns of voiceless fricatives. The fricatives are designed in CVCV contexts in which three peripheral vowels are combined with the fricatives in  $C_1$  and  $C_2$ . Acoustic data, including frequencies and durations of the lower margins, concentration bars, vowel formants and the onset and offset transitions, are measured. It is found that there are three types of coarticulation effects: homorganic, heterorganic and contiguous.

Yan [18] studied the vowel formant pattern and the coarticulation effect in the voiceless stop onset monosyllables, and it is found that there is no effect of tones on formant values of the vowel. However, there is significant effect of aspiration on the formant values at the onset point of the vowel. Chen [19] investigated the intersyllabic anticipatory coarticulation in CVCV sequences, and found that there is no effect of articulation manner on the formant values. There is coarticulatory effect of  $C_2$  on  $V_1$ , as well as  $V_2$  on  $V_1$ . Sun [20] analyzed the coarticulation effect of vowels in read speech, and found that, as the speaking rate increases, the deviation of the vowel formant is also magnified, as is shown in the apparent centralization of the vowel. In fast speech, the formant values of /i/ and /u/ are significantly different, due to the frontness and the backness of the following consonants. Besides, there is also the study on the anticipatory coarticulation in  $V_1\#C_2V_2$  sequences [21], and coarticulatory effect in VCV sequences [22]. However, as far as we know, there have no studies on coarticulatory effect at different vowel points in Chinese.

The present study will investigate the effect of trans-segmental coarticulation in Chinese. In particular, it will try to answer the following questions. Does trans-segmental coarticulation occur in Chinese? What is the extent of coarticulation in Chinese? Coarticulation may be classified as carry-over (left-to-right) or anticipatory (right-to-left) ones, and the present study will focus on carry-over coarticulation.

## II. METHODOLOGY

### A. Speakers and stimuli

The speakers for this experiment were eight native speakers of Standard Chinese, four males and four females. The stimulus list is comprised of 6 stimuli, embedded in three carrier sentences of Chinese, each containing an item from a set of two target words. The target words are 'Biba' and 'Baba', which are supposed to be two persons' names. They are in the form of 'bV<sub>1</sub>ba', with /a/ as the target vowel, and  $V_1$  providing the 'changing' vowel contexts, namely, the contexts of /a/ and /i/.

Labial /b/ is chosen to minimize the effects of consonant articulation on lingual articulations [3]. Within the carrier sentences, the target item is located at the sentence medial position. One example is shown as follow,

Zhe shi Biba de jiejie.  
This is Biba's sister.

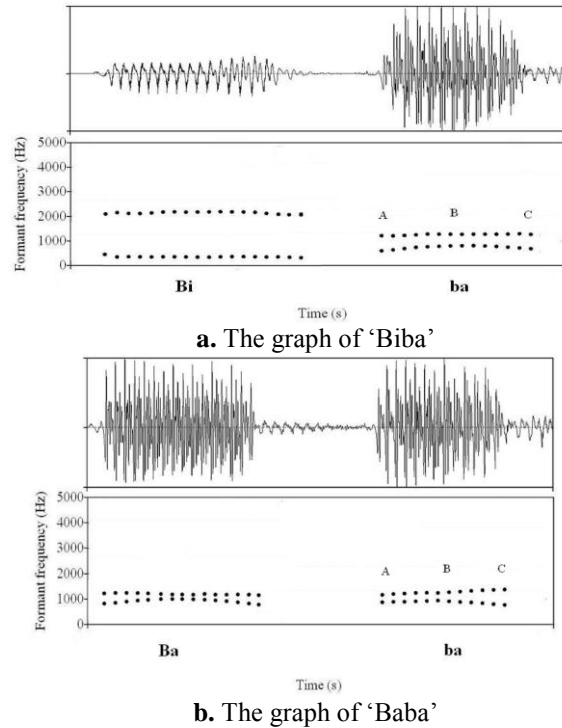
### B. Procedure and measurements

The speakers were asked to read the sentences three times, in random order for each time, in normal speed, so each speaker produced 18 tokens: 6 sentences  $\times$  3 repetitions. In total, 144 tokens were acoustically analyzed (18 tokens  $\times$  8 speakers).

This study aims at investigating the extent of V-to-V coarticulation in VCV sequences, and vowel formants were examined. Formant values were extracted using Praat [23], and the extent of trans-consonantal coarticulation was analyzed by comparing the formant values of the target vowel at three points: the onset point, the middle point and the offset point. That is, formant values at the onset, middle and offset points of the target vowel /a/ were extracted, and the values at different vowel contexts were compared. As is mentioned in the previous subsection, there are two contexts for the target vowel, /i/ and /a/. Coarticulatory effect exists if there is significant difference between the formant values in the two vowel contexts. On the contrary, there is no coarticulatory effect if there is no significant difference. A repeated measures ANOVA was performed for the comparison, and statistic analysis was done in SPSS.

Figure 1 displays the waveforms and formant contours of the two key words 'Biba' (a) and 'Baba' (b). In the graphs, the second syllable 'ba' is the key syllables, and the preceding syllables 'bi' and 'ba' provide the changing vowel contexts. Formant values are extracted from the onset, middle and offset points of the second syllable 'ba', which correspond to point A, B and C in the graphs. Comparison is made for formant values of the target values in the two contexts at each of the three points. Taken the onset point, point A, as an example, since the contexts shown in graph (a) and graph (b) are different, if the formant values at point A in the two contexts are significantly different, it can be concluded that there is an effect of coarticulation at that point. In this study,

comparison will be made for both  $F_1$  and  $F_2$  values at each of the three points, the onset, middle and the offset point.



**Fig. 1:** Waveforms and formant contours of the key words 'Biba' (a) and 'Baba' (b)

### III. RESULT

Figure 2 presents the mean  $F_1$  (a and b) and  $F_2$  (c and d) values of the target vowel /a/ for male (a and c) and female (b and d) speakers, in the contexts of /a/ and /i/, measured at the onset, middle and offset points. One of the aims of this study is to investigate the extent of the coarticulation, so formant is measured at three points, the onset, middle and the offset points of the target vowel. Analysis is given in the following subsections.

#### A. Onset point

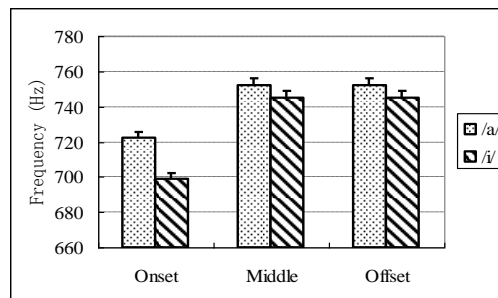
Results from a repeated measures ANOVA show that, when key words occur at sentence medial position, at the onset point of the target vowel, with data of both male and female speakers pooled together, the effect of changing vowel is significant for  $F_2$ , but not for  $F_1$ .  $F_1$ :  $F(1, 71) = 0.04$ ,  $p = 0.847$ ;  $F_2$ :  $F(1, 71) = 47.72$ ,  $p < 0.001$ . That is, coarticulatory effect exists for  $F_2$ , but not for  $F_1$ .

#### B. Middle point

At the middle point of the target vowel, it is shown that, similar to that at the onset point, the effect of changing vowel is significant for  $F_2$ , but not for  $F_1$ .  $F_1$ :  $F(1, 71) = 0.39$ ,  $p = 0.537$ ;  $F_2$ :  $F(1, 71) = 7.28$ ,  $p = 0.009$ . Coarticulatory effect exists for  $F_2$ , but not for  $F_1$ .

#### C. Offset point

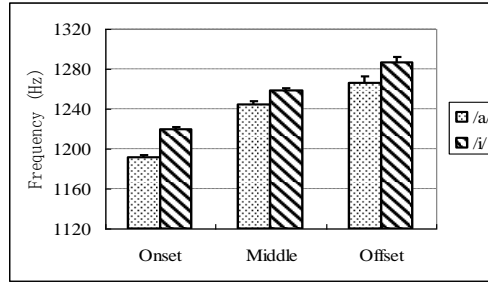
Coming to the offset point of the target vowel, it is shown that the effect of changing vowel is not significant for either  $F_1$  or  $F_2$ .  $F_1$ :  $F(1, 71) = 3.82$ ,  $p = 0.55$ ;  $F_2$ :  $F(1, 71) = 0.25$ ,  $p = 0.621$ . Coarticulatory effect does not exist at that point.



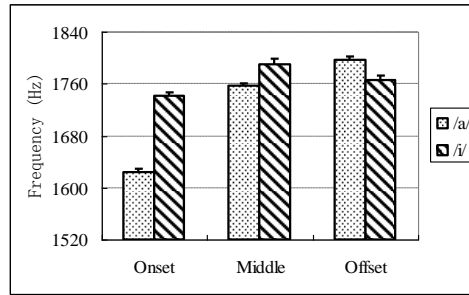
**a.  $F_1$  for male speakers**



b. F<sub>1</sub> for female speakers



c. F<sub>2</sub> for male speakers



d. F<sub>2</sub> for female speakers

**Figure 2:** F<sub>1</sub> (a and b) and F<sub>2</sub> (c and d) values of the target vowel /a/ for male (a and c) and female (b and d) speakers, in the contexts of /a/ and /i/, measured at the onset, middle and offset points

#### IV. DISCUSSION

From the results reported in the previous section it is noted that, trans-segmental coarticulation does exist in Chinese, especially for the second formant value of the vowel. In this experiment, coarticulatory effect is examined at the onset, middle and offset points of the target vowel respectively, and both the first and the second formant values are investigated. It is found that, as far as carry-over coarticulation is concerned, coarticulatory effect exists for the second formant values at the onset and the middle points of the vowel. To be specific, at the onset and the middle points of the vowel, there are significant differences between the two changing vowel contexts for the second formant, but not the first formant.

The effect on the first and the second formants are not consistent with each other at the onset and the middle point of the target vowel, with F<sub>2</sub> affected, while F<sub>1</sub> unaffected. We speculate that the reason for this is that the difference between /a/ and /i/ for F<sub>2</sub> is larger than that for F<sub>1</sub>. According to the report of Bao [24], the mean formant values by 8 male speakers are as follow, /a/: F<sub>1</sub> = 984 Hz, F<sub>2</sub> = 1157 Hz; /i/: F<sub>1</sub> = 283 Hz, F<sub>2</sub> = 2350 Hz. The differences between /a/ and /i/ for F<sub>1</sub> and F<sub>2</sub> are 701 Hz and 1193 Hz respectively. The difference of F<sub>2</sub> is much larger than that of F<sub>1</sub>. When the formant difference is large, the force for triggering the change of course of formant contour is also large. That is to say, vowel coarticulatory effect is more likely to occur on cases with great formant difference, therefore, it occurs on F<sub>2</sub>, not on F<sub>1</sub>.

When the offset point of the target vowel is investigated, it is found that coarticulatory effect does not exist. That is, for either F<sub>1</sub> or F<sub>2</sub>, there is no coarticulatory effect. This result is caused by the ‘distance effect’: at the onset and the middle points, the distance from the measured point to the changing vowels is close, and the effect exists; while at the offset point, the distance gets farther, and the effect disappears.

Magen [15] investigated the extent of vowel-to-vowel coarticulation in English trisyllabic utterances, and it was found that coarticulatory effects can, in some instances, extend beyond the bounds that previous research had assumed; coarticulatory effect can extend from one full vowel, through the medial schwa, and into the midpoint of the next full vowel. He proposed that foot does not define the domain over which coarticulatory effects operate. However, in the present study, it is found that coarticulatory effect does not extend to the end of the vowel. We speculate this is because that Chinese is of different language typology from English.

One of the well-known properties of language typology is the rhythm unit of a language. Languages have been categorized as mora-timed, such as Japanese, stress-timed, like English and German, and syllable-timed, such as Chinese and French [25-26]. English is a stress-timed language, and the unstressed syllables are quite weak, so it is possible for coarticulatory effect to extend to the third syllable in English. Chinese is a syllable-timed language, in which syllables are rarely as weak as the unstressed ones in English. Compared to English, the degree of articulatory constraint (DAC) in Chinese is high. Therefore, the coarticulatory effect in Chinese is not as great as that in English.

## V. CONCLUSION

In this study, the vowel-to-vowel coarticulation effect in bi-syllabic words in Chinese is analyzed, and it is found that trans-segmental coarticulation exists in Chinese, especially for the second formant value of the vowel. Coarticulation is more likely to occur on  $F_2$  because the difference between /a/ and /i/ of  $F_2$  is larger than that of  $F_1$ . Because of the ‘distance effect’, coarticulatory effect exists at the onset and middle points, and disappeared at the offset point of the target vowel. In Chinese, coarticulatory effect is not as great as that in English, as the two languages are of different prosodic typology.

This study is significant for speech engineering. In speech synthesis, the effect of trans-segmental coarticulation must be taken into consideration, especially for the second formant value of the vowel. The first formant is not affected, so it is not necessary to mind about the effect on  $F_1$ . Nor is it necessary to consider the trans-syllabic coarticulatory effect at the end of the vowel, as in Chinese, coarticulatory effect does not extend to the end of the syllable. Therefore, this study is helpful for speech engineering technology.

## ACKNOWLEDGMENT

This work was supported in part by the Social Science Research Project of Guangdong Province, Grant No. GD11CWW04, as well as the Humanity and Social Science Research Project for Colleges in Guangdong Province, Grand No. 11WYXM012.

## REFERENCES

- [1] W. F. Katz, “Anticipatory coarticulation and aphasia implications for phonetic theories”, *Journal of Phonetics*, (28), pp. 313–334, 2000.
- [2] S. Y. Manuel, “Cross-language studies: relating language-particular coarticulation patterns to other language-particular facts”, in W. J. Hardcastle and N. Hewlett (eds), *Coarticulation: theory, data and techniques*, Cambridge University Press, Cambridge, U.K., pp. 179–198, 1999.
- [3] S. E. G. Öhman, “Coarticulation in VCV utterances: spectrographic measurements,” *Journal of the Acoustical Society of America*, Vol. 39, pp. 151–168, 1966.
- [4] P. S. Beddor, J. D. Harnsberger and S. Lindemann, “Language-specific patterns of vowel-to-vowel coarticulation: acoustic structures and their perceptual correlates,” *Journal of Phonetics* (30), pp. 591–627, 2002.
- [5] S.Y. Manuel and R. A. Krakow, “Universal and language particular aspects of vowel-to-vowel coarticulation,” *Haskins Laboratories Status Reports on Speech Research*, SR-77/78, pp. 69–78, 1984.
- [6] T. Cho, “Prosodically conditioned strengthening and vowel-to-vowel coarticulation in English,” *Journal of Phonetics*, (32), pp. 141–176, 2004.
- [7] D. Recasens, “Vowel-to-vowel coarticulation in Catalan VCV sequences,” *Journal of the Acoustical Society of America*, 76, pp. 1624–1635, 1984.
- [8] D. Recasens, “An acoustic analysis of V-to-C and V-to-V coarticulatory effects in Catalan and Spanish VCV sequences,” *Journal of Phonetics*, 15, pp. 299–312, 1987.
- [9] S. Y. Manuel, “The role of contrast in limiting vowel-to-vowel coarticulation in different languages,” *Journal of the Acoustical Society of America*, 88, pp. 1286–1298, 1990.
- [10] D. H. Whalen, “Coarticulation is largely planned,” *Journal of Phonetics*, 18, pp.3–35, 1990.
- [11] E. T. Purcell, “Formant frequency patterns in Russian VCV utterances,” *Journal of the Acoustical Society of America*, 66, pp. 1691–1702, 1979.
- [12] T. Gay, “Articulatory movements in VCV sequences,” *Journal of the Acoustical Society of America*, (62), pp. 183–193, 1977.

- [13] P. West, "The extent of coarticulation of English liquids: An acoustic and articulatory study," In *Proceedings of the international conference of phonetic sciences*, San Francisco, pp. 1901–1904, 1999.
- [14] S. Heid and S. Hawkins, "An acoustical study of long-domain /r/ and /l/ coarticulation," In *Proceedings of the fifth seminar on speech production: Models and data*, Bavaria, Germany: Kloster Seeon, pp. 77–80, 2000.
- [15] H. S. Magen, "The extent of vowel-to-vowel coarticulation in English," *Journal of Phonetics*, (25), pp. 187–205, 1997.
- [16] M. Grosvald, "Interspeaker variation in the extent and perception of long-distance vowel-to-vowel coarticulation," *Journal of Phonetics*, (37), pp. 173–188, 2009.
- [17] Z. Wu and G. Sun, "Acoustic coarticulatory patterns of voiceless fricatives in CVCV in Standard Chinese", *Report on Phonetic Research*, Institute of Linguistics, CASS, pp. 1–19, 1990.
- [18] J. Yan, "A study of the vowel formant pattern and the coarticulation in the voiceless stop initial monosyllable of Standard Chinese", *Report on Phonetic Research*, Institute of Linguistics, CASS, pp. 30–54, 1990.
- [19] X. Chen, "A study on the coarticulation of segments in Chinese, Chinese Linguistics", *Zhongguo Yuwen*, Vol. 5, pp. 345–350, 1997.
- [20] G. Sun, "Vowel segmental coarticulation in read speech in Standard Chinese", *Report on Phonetic Research*, Institute of Linguistics, Chinese Academy of Social Sciences, 1998.
- [21] Y. Li and J. Kong, "Anticipatory coarticulation in V1#C2V2 sequences in standard Chinese," *Journal of Tsinghua University*, Vol. 51, No. 9, pp 1220-1225, 2011.
- [22] M. Wang, W. Yan and Z. Xiong, "Coarticulatory effect in VCV sequences in Chinese," *Journal of Tsinghua University*, Vol. 51, No. 9, pp 1244-1248, 2011.
- [23] P. Boersma, "Praat, a system for doing phonetics by computer", *Glott International*, 5:9/10, pp. 341–345, 2001.
- [24] H. Bao, The physiological explanation of monophthongs in Chinese, *Zhongguo Yuwen* (2): pp. 117-127, 1984.
- [25] K. L. Pike, *The intonation of American English*. Ann Arbor: University of Michigan Press, 1945.
- [26] I. Lehiste, *Suprasegmentals*, Cambridge, MA: MIT Press, 1970.